Normalized Log-Likelihood Particle ID from MIPP PID detectors

Holger Meyer

September 11, 2009

Abstract

This document summarizes the algorithms for normalized PID LLs from the four MIPP PID detectors. It does not cover the global PID.

1 Introduction

The four particle identification detectors in MIPP (TPC dE/dx, ToF, Ckov, and RICH, ignoring ECAL and HCAL here) identify charged particles of all momenta. Due to finite resolutions in TPC dE/dx and ToF timing and Poisson fluctuations in Ckov and RICH the particle ID does not result in a unique ID with 100% certainty. Thus the information from each PID detector is best represented in the form of likelihoods.

1.1 Likelihood basics

This section briefly reviews the concept of likelihood. The web contains a lot of good resources for those who need to refresh their knowledge [1, 2].

A likelihood describes the probablility of an observation (e.g. tof=9.2 ns) being due to an unknown parameter (e.g. PID of the track). The likelihood function is not unique and scaling it by a constant does not change the result. The likelihoods of different PIDs are compared in likelihood ratios and the constant factor cancels in the ratio. As is common we use the natural logarithm of the likelihoods instead of the likelihoods themselves. The ratios of likelihoods then become differences in log-likelihoods.

In order to be able to combine the log-likelihoods from different detectors we want to provide *normalized* log-likelihoods for each detector.

2 TPC

The particle identification for the TPC is provided by measuring average dE/dx for a track, q_{meas}. The expected (or predicted) dE/dx is given by the Bethe-Bloch equation (and by an empirical variation for the electron dE/dx) for each particle ID as $q_{pred}(PID)$. The observation deviates from the predicted values due to a resolution σ.

If we were to measure (at a fixed momentum) the dE/dx of many particles of type PID we would obtain a Gaussian distribution centered at $q_{pred}(PID)$ with a width of σ. Thus the probability of observing a certain q_{meas} is given by

$$P(q_{meas}|q_{pred}(PID),\sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(p_{meas}-ppred)^2}{2\sigma^2}} \tag{1}$$

The factor of $\frac{1}{\sigma\sqrt{2\pi}}$ normalizes the probability to unity. (The probability of measuring any value of q_{meas} is 100%.)

The likelihood of a certain $q_{pred}(PID)$ being the cause of the observation q_{meas} is given by the same equation:

$$L \equiv L(q_{pred}(PID)|q_{meas},\sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(p_{meas}-p_{pred})^2}{2\sigma^2}} \tag{2}$$

We now take the natural logarithm and simplify:

$$\begin{aligned}
LL \equiv \ln L &= \ln \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(p_{meas}-p_{pred})^2}{2\sigma^2}} && (3)\\
&= \left(\ln \frac{1}{\sqrt{2\pi\sigma^2}}\right) + (-0.5)\left(\frac{p_{meas}-p_{pred}}{\sigma}\right)^2 && (4)\\
&= -0.5\left(\ln 2\pi\sigma^2\right) + (-0.5)\left(\frac{p_{meas}-p_{pred}}{\sigma}\right)^2 && (5)\\
&= -0.5\left(\left(\frac{p_{meas}-p_{pred}}{\sigma}\right)^2 + \ln\left(2\pi\sigma\sigma\right)\right) && (6)
\end{aligned}$$

The last line presents the calculation as it is done in the `TPCRPID` module.

3 TOF

All considerations for the TPC directly apply for the TOF where q now represents the time of flight for the track and σ is the time of flight resolution.

In the calculation in the `TOFTReco` module the log likelihoods were calculated with the normalization wrong by a factor of $\sqrt{2\pi}$. This has been corrected in the tagged release `R09.09.10` and later.

4 Ckov

For the multicell Cherenkov detector we compute the expected number of photons for each particle hypothesis and determine the number of observed photons and resolution from the data. The situation is complicated by the fact that tracks can overlap such that light observed in a given photomultiplier tube can be due to one track or another (or both). It should be noted that a track passing through the Ckov volume with no light observed in the corresponding set of pmts gathers information on its identity

from this absence of light. The likelihood is calculated separately for the cases of light being observed and light not being observed.

The algorithm is implemented in the `GetLLs` method of `CkovMatchTrk`. It will be summarized here in a later revision of this document.

5 RICH

In the RICH light is observed on an array of photomultiplier tubes. The light in a particular pmt may be due to the ring of any of the tracks or a combination of several of them.

The RICH LL algorithm will be summarized here in a later version of this document.

6 Summary

We describe the log likelihood algorithms for the four MIPP PID detectors. The likelihoods are properly normalized so that a global pid algorithm can combine these individual LLs and the a priory particle fractions at each momentum into a global PID LL.

References

[1] Wikipedia entry `http://en.wikipedia.org/wiki/Likelihood_function` and other related entries.

[2] Weisstein, Eric W. "Likelihood." From MathWorld–A Wolfram Web Resource. `http://mathworld.wolfram.com/Likelihood.html` and pages linked from there.